Hydraulic and transport parameter assessment using column infiltration experiments

A. Younes[1,2,3], T.A. Mara[4], M. Fahs[1], O. Grunberger[2], Ph. Ackerer[*,1]

[1] LHyGES, Université de Strasbourg/ENGEES, CNRS, 1 rue Blessig, 67084 Strasbourg, France.
[2] IRD UMR LISAH, F-92761 Montpellier, France.
[3] LMHE, Ecole Nationale d'Ingénieurs de Tunis, Tunisie
[3] Université de La Réunion, PIMENT, 15 Avenue René Cassin, BP 7151, 97715 Moufia, La Réunion.

* Contact person: Ph. Ackerer
E-mail: ackerer@unistra.fr

Abstract

The quality of statistical calibration of hydraulic and transport soil properties is studied for infiltration experiments in which, over a given period, tracer-contaminated water is injected into an hypothetical column filled with a homogeneous soil. The saturated hydraulic conductivity, the saturated and residual water contents, the Mualem-van Genuchten shape parameters and the longitudinal dispersivity are estimated in a Bayesian framework using the Markov Chain Monte Carlo (MCMC) sampler. The impact of the kind of measurement sets (water content, pressure inside the column, cumulative outflow and outlet solute concentration) and that of the solute injection duration is investigated by analyzing the calibrated model parameters and their confidence intervals for different scenarios. The results show that the injection period has a significant effect on the quality of the estimation, in particular, on the posterior uncertainty range of the parameters. All hydraulic and transport parameters of the investigated soil can be well estimated from the experiment using only the outlet concentration and cumulative outflow, which are measured non-intrusively. An improvement of the identifiability of the hydraulic parameters is observed when the pressure data from measurements taken inside the column are also considered in the inversion.

Keywords

Infiltration experiment, Richards' equation, Statistical calibration, Markov Chain Monte Carlo, Uncertainty ranges.

1. Introduction

The soil parameters that influence water flow and contaminant transport in unsaturated zones are not generally known *a priori* and have to be estimated by fitting model responses to observed data. The unsaturated soil hydraulic parameters can be (more or less accurately) estimated from dynamic flow experiments (*e.g.*, Hopmans et al., 2002; Vrugt et al., 2003a; Durner and Iden, 2011; Younes et al., 2013). Several authors have investigated different types of transient experiments and boundary conditions suited for a reliable estimation of soil hydraulic properties (e.g. van Dam et al., 1994; Simunek and van Genuchten, 1997; Inoue et al, 1998; Durner et al, 1999). Soil hydraulic properties are often estimated using inversion of one-step (Kool et al., 1985; van Dam et al., 1992) or multistep (Eching et al., 1994; van Dam et al., 1994) outflow experiments or controlled infiltration experiments (Hudson et al., 1996).

Kool et al. (1985) and Kool and Parker (1988) suggested that the transient experiments should cover a wide range in water contents to obtain a reliable estimation of the parameters. Van Dam et al. (1994) have shown that more reliable parameter estimates are obtained by increasing the pneumatic pressure in several steps instead of a single step. The multistep outflow experiments are the most popular laboratory methods (e.g., Eching and Hopmans, 1993; Eching et al., 1994; van Dam et al., 1994; Hopmans et al., 2002). However, their application is limited by expensive measurement equipment (Nasta et al., 2011).

Infiltration experiments have been investigated by Mishra and Parker (1989) to study the reliability of hydraulic and transport estimated parameters for a soil column of 200 cm using measurements of water content, concentration and water pressure inside the column. They showed that the simultaneous estimation of hydraulic and transport properties yields to smaller estimation errors for model parameters than the sequential inversion of hydraulic properties from the water content and/or pressure head followed by the inversion of transport properties from concentration data (Mishra and Parker, 1989).

Inoue et al. (2000) performed infiltration experiments using a soil column of 30 cm. Pressure head and solute concentration were measured at different locations. A constant infiltration rate was applied to the soil surface and a balance was used to measure the cumulative outflow. They showed that both hydraulic and transport parameters can be assessed by the combination of flow and transport experiments.

Furthermore, infiltration experiments were often conducted in lysimeters for pesticide leaching studies. Indeed, lysimeter experiments are generally used to assess the leaching risks of pesticides using soil columns of around 1.2 m depth which is the standard scale for these types of experiments (Mertens et al, 2009; Kahl et al., 2015). Before performing the column leaching experiment, several infiltration-outflow experiments are often realized to estimate the soil hydraulic parameters (Kahl et al., 2015; Dusek et al, 2015).

The key objective of the present study is to evaluate the reliability of different experimental protocols for estimating hydraulic and transport parameters and their associated uncertainties for column experiments. We consider the flow and the transport of an inert solute injected into a hypothetical column filled with a homogeneous sandy clay loam soil. We assume that flow can be modelled by the Richards' equation (RE) and that the solute transport can be simulated by the classical advection-dispersion model. Furthermore, the Mualem and van Genuchten (MvG) models (Mualem 1976, van Genuchten 1980) are chosen to describe the retention curve and to relate the hydraulic conductivity of the unsaturated soil to the water content. The estimation of the flow and transport parameters through flow-transport model inversion is investigated for two injection periods of the solute and different data measurement scenarios.

Inverse modelling is often performed using local search algorithms such as the Levenberg-Marquardt algorithm (Marquardt, 1963). The later is computationally efficient to evaluate the optimal parameter set (Gallagher and Doherty, 2007). Besides, the degree of uncertainty in the estimated parameters, expressed by their confidence intervals, is often calculated using a first-order approximation of the model near its minimum (Carrera and Neuman, 1986, Kool and parker, 1988). However, as stated by Vrugt and Bouten (2002), parameter interdependence and model nonlinearity occurring in hydrologic models may violate the use of this first approximation to obtain accurate confidence intervals of each parameter. Therefore, in this work, the estimation of hydraulic and transport parameters is performed in a Bayesian framework using the Markov Chain Monte Carlo (MCMC) sampler (Vrugt and Bouten, 2002; Vrugt et al., 2008). Unlike classical parameter optimization algorithms, the MCMC approach generates sets of parameter values randomly sampled from the posterior joint probability distributions, which are useful to assess the quality of the estimation. The MCMC samples can be used to summarize parameter uncertainties and to perform predictive uncertainty (Ades and Lu, 2003).

Hypothetical infiltration experiments are considered for a column of 120 cm depth, initially under hydrostatic conditions, free of solute and filled with a homogeneous sandy clay loam soil. Continuous flow and solute injection are performed during a time period T_{inj} at the top of the column and with a zero pressure head at the bottom. The unknown parameters for the water flow are the hydraulic parameters: k_s [L.T^{-1}], the saturated hydraulic conductivity; θ_s [L^3.L^{-3}], the saturated water content; θ_r [L^3.L^{-3}], the residual water content; and α [L^{-1}] and n [−], the MvG shape parameters. The only unknown parameter of the tracer transport is the longitudinal dispersivity, a_L [L].

Several scenarios corresponding to different sets of measurements are investigated to address the following questions:

1) Can we obtain an appropriate estimation of all flow and transport parameters from tracer-infiltration experiments, even though a limited range in water content is covered

(only moderately dry conditions are obtained because of gravity drainage conditions prescribed at the bottom of the soil column)?

2) What is the optimal set of measurements for the estimation of all the parameters? Can we use only non-intrusive measurements (cumulative outflow and concentration breakthrough curve) or are intrusive measurements of pressure heads and/or water contents inside the column unavoidable?

3) Is there an optimal design for the tracer injection?

For this purpose, synthetic scenarios are considered in the sequel in which data from numerical simulations are used to avoid the uncontrolled noise of experiments that could bias the conclusions.

The paper is organized as follows. The mathematical models describing flow and transport in the unsaturated zone are detailed in section 2. Section 3 describes the MCMC Bayesian parameter estimation procedure used in the DREAM$_{(ZS)}$ sampler. Section 4 presents the different investigated scenarios and discusses the results of the calibration in terms of mean parameter values and uncertainty ranges for each scenario. Conclusions are given in section 5.

2. Unsaturated flow-transport model

We consider a uniform soil profile in the column and an injection of a solute tracer such as bromide, as described in Mertens et al. (2009). The unsaturated water flow in the vertical soil column is modeled with the one-dimensional pressure head form of the RE:

$$\begin{cases} \left(c(h) + S_s \dfrac{\theta}{\theta_S} \right) \dfrac{\partial h}{\partial t} = \dfrac{\partial q}{\partial z} \\ q = K(h) \left(\dfrac{\partial h}{\partial z} - 1 \right) \end{cases}, \qquad (1)$$

where h [L] is the pressure head; q [L.T^{-1}] is the Darcy velocity; z [L] is the depth, measured as positive in the downward direction; S_s (-) is the specific storage; θ and θ_s [L^3.L^{-3}] are the actual and saturated water contents, respectively; $c(h)$ [L^{-1}] is the specific moisture capacity; and $K(h)$[L.T^{-1}] is the hydraulic conductivity. The latter two parameters are both functions of the pressure head. In this study, the relations between the pressure head, conductivity and water content are described by the following standard models of Mualem (1976) and van Genuchten (1980):

$$S_e(h) = \frac{\theta(h)-\theta_r}{\theta_s-\theta_r} = \begin{cases} \dfrac{1}{\left(1+|\alpha h|^n\right)^m} & h<0 \\ 1 & h \geq 0 \end{cases}, \qquad K(S_e) = K_s S_e^{1/2}\left[1-\left(1-S_e^{1/m}\right)^m\right]^2 \tag{2}$$

where S_e (-) is the effective saturation, θ_r [L^3.L^{-3}] is the residual water content, K_s [L.T^{-1}] is the saturated hydraulic conductivity, and $m = 1-1/n$, α [L^{-1}] and n (-) are the MvG shape parameters.

The tracer transport is governed by the following convection-dispersion equation:

$$\frac{\partial(\theta C)}{\partial t} + \frac{\partial(qC)}{\partial z} - \frac{\partial}{\partial z}\left(\theta D \frac{\partial C}{\partial z}\right) = 0 \tag{3}$$

where C [M.L^{-3}] is the concentration of the tracer, D [L^2.T^{-1}] is the dispersion coefficient in which $D = a_l\, q + d_m$ and a_l [L] is the dispersivity coefficient of the soil and d_m [L^2.T^{-1}] is the molecular diffusion coefficient, which is set as 1.04 10^{-4} cm^2/min.

The transport equation (3) is coupled with the flow equation (1) by the water content θ and the Darcy's velocity q. The initial conditions are as follows: a hydrostatic pressure distribution with zero pressure head at the bottom of the column $(z = L)$ and a solute concentration of zero inside the whole column. An infiltration with a flux q_{inj} of contaminated water with a concentration C_{inj} is then applied at the upper boundary condition (z = 0) during a period T_{inj}. Hence, the boundary conditions at the top of the column can be expressed as:

$$for\ 0 < t \le T_{inj} \begin{cases} K\left(\dfrac{\partial h}{\partial z} - 1\right) = q_{inj} \\ \theta D \dfrac{\partial C}{\partial z} + qC = q_{inj} C_{inj} \end{cases} \qquad for\ t > T_{inj} \begin{cases} K\left(\dfrac{\partial h}{\partial z} - 1\right) = 0 \\ C_{inj} = 0 \end{cases}, \qquad (4)$$

A zero pressure head is maintained at the lower boundary $(z = L)$ of the column and a zero concentration gradient is used as the lower boundary condition for the solute transport, namely,

$$(h)_{z=l} = 0 \qquad \left(\frac{\partial C}{\partial z}\right)_{z=l} = 0 \qquad (5)$$

In the sequel, the infiltration rate and the injected solute concentration are $q_{inj} = 0.015$ cm/min and $C_{inj} = 1$ g/cm^3, respectively. The system (1)-(5) is solved using the standard finite difference method for both flow and transport spatial discretization. A uniform mesh of 600 cells is employed. Temporal discretization is performed with the high-order method of lines (MOL) (e.g., Miller et al., 1998; Tocci et al., 1997; Younes et al., 2009; Fahs et al., 20011). Error checking, robustness, order selection and adaptive time step features, available in sophisticated solvers, are applied to the time integration of partial differential equations (Tocci et al., 1997). The MOL has been successfully used to solve RE in many studies (e.g., Farthing et al., 2003; Miller et al., 2006; Li et al., 2007; Fahs et al., 2009). Details on the use of the MOL for solving RE are described in Fahs et al. (2009).

3. Bayesian parameter estimation

The vector of unknown parameters that has to be identified by model calibration is $\xi = (k_s, \theta_s, \theta_r, \alpha, n, a_L)$. To analyze the performance of the model calibration procedures, a reference solution is generated by simulating the flow-transport problem (1)-(5) using the following parameter values (corresponding to a sandy clay loam soil): $k_s = 50\,cm/day$, $\theta_s = 0.43$, $\theta_r = 0.09$, $\alpha = 0.04\,cm^{-1}$, $n = 1.4$ and $a_l = 0.2\,cm$. Four types of variables are extracted from the results of the simulation: the pressure head and water content 5 cm below the top of the column, the cumulative outflow and the solute breakthrough concentration at the outflow of the column. These four data series are modified by adding a normally distributed white noise using the following standard deviations: $\sigma_h = 1\,cm$ for the pressure head, $\sigma_\theta = 0.02$ for the water content, $\sigma_Q = 0.1$ cm for the cumulative outflow and $\sigma_C = 0.01$ g/cm^3 for the exit concentration. These perturbations mimic measurement errors and the resulting values of water pressure, water content, cumulative outflow and solute breakthrough concentration are considered as measurements in the following.

The flow-transport model is used to analyze the effects of different measurement sets on parameter identification. For this purpose, we adopt a Bayesian approach that involves the parameter joint posterior distribution (Vrugt et al., 2008). The latter is assessed with the DREAM$_{(ZS)}$ MCMC sampler (Laloy and Vrugt, 2012). This software generates random sequences of parameter sets that asymptotically converge toward the target joint posterior distribution (Gelman et al., 1997). Thus, if the number of runs is sufficiently high, the generated samples can be used to estimate the statistical measures of the posterior distribution, such as the mean and variance among other measures.

The Bayes theorem states that the probability density function of the model parameters conditioned onto data can be expressed as:

$$p\left(\xi \mid \boldsymbol{y}_{mes}\right) \propto p\left(\boldsymbol{y}_{mes} \mid \xi\right) p\left(\xi\right), \tag{6}$$

where $p\left(\xi \mid \boldsymbol{y}_{mes}\right)$ is the likelihood function measuring how well the model fits the observations $\boldsymbol{y}_{mes}$, and $p\left(\xi\right)$ is the prior information about the parameter before the observations are made. Independent uniform priors within the ranges reported in Table 1 are chosen. In this work, a Gaussian distribution defines the likelihood function because the *observations* are simulated and corrupted with Gaussian errors. Hence, the parameter posterior distribution is expressed as:

$$p\left(\xi \mid \boldsymbol{y}_{mes}\right) \propto exp\left(-\frac{SS_h\left(\xi\right)}{2\sigma_h^2}-\frac{SS_\theta\left(\xi\right)}{2\sigma_\theta^2}-\frac{SS_Q\left(\xi\right)}{2\sigma_Q^2}-\frac{SS_C\left(\xi\right)}{2\sigma_C^2}\right), \tag{7}$$

where $SS_h\left(\xi\right)$, $SS_\theta\left(\xi\right)$, $SS_Q\left(\xi\right)$ and $SS_C\left(\xi\right)$ are the sums of the squared differences between the observed and modeled data of the pressure head, water content, cumulative outflow and output concentration, respectively. For instance, $SS_h\left(\xi\right)=\sum_{k=1}^{Nh}\left(h_{mes}^{(k)}-h_{mod}^{(k)}\left(\xi\right)\right)^2$, which includes the observed $h_{mes}^{(k)}$ and predicted $h_{mod}^{(k)}$ pressure heads at time t_k for the number of pressure head observations Nh.

Bayesian parameter estimation is performed hereafter with the $\mathrm{DREAM}_{(ZS)}$ software (Laloy and Vrugt, 2012), which is an efficient MCMC sampler. $\mathrm{DREAM}_{(ZS)}$ computes multiple sub-chains in parallel to thoroughly explore the parameter space. Archives of the states of the sub-chains are stored and used to allow a strong reduction of the "burn-in" period in which the sampler generates individuals with poor performances. Taking the last 25% of individuals of the MCMC (when the chains have converged) yields multiple sets of parameters, ξ, that adequately fit the model onto observations. These sets are then used to estimate the updated parameter distributions, the pairwise parameter correlations and the uncertainty of the model predictions. As suggested in Vrugt et al. (2003b), we consider that the posterior distribution is stationary if the Gelman and Ruban (1992) criterion is ≤ 1.2 .

4. Results and discussion

In this section, the identifiability of the parameters is investigated for 7 different scenarios of measurement sets (Table 1). In the first scenario, only measured pressure heads and cumulative outflow are used for the calibration. The scenarios 2 to 5 investigate the benefit of adding measured water contents and/or solute outlet concentrations to pressure heads and outflow. The last scenarios (6, 7) investigate the use of measured cumulative outflow and concentration breakthrough at the column outflow because these measurements do not require intrusive techniques. Scenarios 5 to 7 investigate as well the effects of solute injection duration on the identifiability of the parameters.

In all cases, the MCMC sampler was run with 3 simultaneous chains for a total number of 50000 runs. Depending on the scenario, the MCMC required between 5000 and 20000 model runs to reach convergence and was terminated after 30000 runs. The last 25% of the runs that adequately fit the model onto observations are used to estimate the updated probability density function (pdf).

4.1. The data sets for parameter estimation

The data sets obtained from solving the flow-transport problems (1)-(5) using the parameters given in section 2 are shown in Fig. 1. The pressure head at 5 cm from the top of the column (Fig. 1a) increases from its initial hydrostatic negative value (-115 cm) and reaches a plateau (-1.75 cm) in less than 100 minutes during the injection period. After the injection is finished, it progressively decreases due to the drainage caused by the gravity effect. A similar behavior is observed for the water content at the same location (Fig. 1b), where the value of the plateau is close to the saturation value. The cumulative outflow (Fig. 1c) starts to increase at approximately 1000 min after the beginning of the injection. It shows an almost linear behavior until 5500 min. It then slowly increases with an asymptotic behavior due to the natural drainage after the end of the injection period. Fig. 1d displays the water saturation as a function of the pressure head. It is worth noting that only a few part of this curve is described during the infiltration experiment. Indeed, only moderate dry conditions are established because the minimum pressure head reached in the column is -120 cm, which corresponds to the initial pressure head at the top of the column.

The breakthrough concentration curve (Fig. 1e) shows a sharp front, which starts shortly after 3000 min. Note that if the injection of both water and contaminant are stopped once the solute reaches the output. For an injection period of 3000 min, the breakthrough curve exhibits a smoother progression (Fig. 1f).

The data considered as measurements, which are used as conditioning information for model calibration, are also shown in Fig. 1. In Fig. 1b, the water content seems to be more affected by the perturbation of data than the pressure head and cumulative outflow. This phenomenon is due to the relative importance of the measurement errors of the water content often observed with time-domain-reflectometry probes and to the weak variations of the water content during the infiltration experiment. The perturbation of the breakthrough curve is relatively small because of the low added noise since output concentrations can be accurately measured. The perturbations of the pressure head and cumulative outflow seem weak because of the large variation of these variables during the experiment.

4.2. Results of the parameter estimation

The uncertainty model parameters are assumed to be distributed uniformly over the ranges reported in Table 1. This table also lists the reference values used to generate data observations before perturbation. Seven scenarios are considered, corresponding to different sets of measurements for the estimation of the hydraulic and transport soil parameters (Table 2).

The MCMC results of the seven studied scenarios are given in Figs. 2 to 8. The "on-diagonal" plots in these figures display the inferred parameter distributions, whereas the "off-diagonal" plots represent the pairwise correlations in the MCMC sample. If the draws are independent, non-sloping scatterplots should be observed. However, if a good value of a given parameter is conditioned by the value of another parameter, then their pairwise scatterplot should show a narrow sloping stripe. The sensitivity of parameters is obtained by comparing prior to posterior parameter distribution. A significant difference between the two distributions for a parameter indicates high model sensitivity to that parameter (Dusek et al., 2015).

To facilitate the comparison between the different scenarios, Figs. 9 to 14 show the mean and the 95% confidence intervals of the final MCMC sample that adequately fit the model onto observations for each scenario, and Table 3 summarizes the pairwise parameter correlations.

Fig. 2 shows the inferred distributions of the parameters identified with the MCMC sampler using only the pressure and cumulative outflow measurements (scenario 1). The parameters k_s, α and n are well estimated; their prior intervals of variation are strongly narrowed and they essentially show bell-shaped posterior distributions. This shows the high sensitivity of the model responses to these parameters.

The parameter k_s is strongly correlated to α (0.94) and n (-0.97). These results confirmed the results of Eching and Hopmans (1994) on multistep outflow experiments who found that the inverse solution technique is greatly improved when both cumulative outflow and pressure head data from some positions inside the column are used. The two water contents related parameters are strongly correlated (0.96) and cannot be identified accurately because the water retention relationship depends on the difference between θ_s and θ_r and only this difference is identifiable. Note that the prior intervals of θ_r and θ_s which are respectively $[0.05, 0.2]$ and $[0.3, 0.5]$ have changed to the posterior intervals $[0.05, 0.16]$ and $[0.39, 0.5]$ because the target difference should be $\theta_s - \theta_r = 0.34$. In the literature, van Genuchten and Nielsen (1985), Eching and Hopmans (1993) and Zurmühl (1996) considered that saturated water content is determined independently and considered only θ_r to be an empirical parameter that should be fitted to the data.

The dispersivity coefficient a_l has not been identified in this first scenario.

The MCMC results in Fig. 3 show that water content measurements throughout the experiment (scenario 2) allow the estimation of both the residual and saturated water contents. The parameter θ_r strongly correlates to k_s (-0.94) and n (0.98) and the parameter k_s remains strongly related to α (0.94) and n (-0.98). Although the water content data are subject to relatively high measurement errors, a good estimation is obtained for θ_s and θ_r. The parameters k_s, α and n are estimated with the same accuracy as for the first scenario. All parameters (except the dispersivity coefficient) are highly sensitive since their posterior intervals of variations are strongly reduced compared to the prior intervals. Moreover, the prior uniform distributions give place to almost Gaussian posterior distributions. These results show that, although Kool et al. (1985) and Kool and Parker (1988) suggested that the transient experiments should cover a wide range in water content, an appropriate estimation of all parameters can be obtained with the infiltration experiment even though a limited range in water content is covered.

When the concentration measurements are also considered in the inversion (scenario 3), the results depicted in Fig. 4 show very significant correlations between k_s and θ_r (-0.94), k_s and α (0.91), k_s and n (-0.97) and n and θ_r (0.99). The posterior uncertainty ranges of k_s,

α, n and θ_r are similar to the previous scenarios. Those of θ_s and a_l are strongly reduced, leading to a good identification of these parameters when using C measurements (Figs. 10 and 14). A better estimate of the saturated water content is obtained because advective transport is a function of this variable.

In the inversion procedure of scenario 4, the measurements of the water content are not considered. This scenario leads to the same quality of the estimation for the parameters k_s, θ_r, α and n (Figs. 9, 11, 12, 13) and similar correlations between the parameters as in the previous scenario. This result shows that the intrusive water content measurements, which are subject to more significant measurement errors than the output concentration, are not required if the output concentration is measured. Compared with the results of scenario 2, it can be concluded that better parameter estimations are obtained using h, Q and C data than using h, Q and θ data, especially for θ_s. Therefore, using C instead of θ measurements in combination with h and Q measurements allows the estimation of a_l and yields better estimate of θ_s.

The pressure head, cumulative outflow and concentration measurements are used in the estimation procedure of scenario 5, but the injection period is now reduced to $T_{inj} = 3000\,\text{min}$. The obtained results (Fig. 6) show the same correlations between the parameters as for $T_{inj} = 5000\,\text{min}$. For the parameters k_s, θ_s, θ_r, α and n, almost the same mean estimates are obtained as for scenario 4. However, the parameters are better identified (Figs. 9 to 13). Indeed, the uncertainty of these parameters is smaller because the credible interval is reduced by a factor of 25% for k_s, 8% for θ_s, 26% for θ_r, 10% for α and 25% for n when compared to the results obtained using $T_{inj} = 5000\,\text{min}$. The parameter a_l is also much better estimated than in the previous scenario. Its mean value approaches the reference solution and the posterior uncertainty range is reduced by approximately 75% (Fig. 14).

In scenario 6, the pressure head measurements are removed and only non-intrusive measurements (Q and C data) are used for the calibration with an injection period of $T_{inj} = 5000\,\mathrm{min}$. These kind of nonintrusive measures have been used by Mertens et al. (2009) to estimate some of hydraulic and pesticides leaching parameters. The results depicted in Fig. 7 show high correlations only between k_s and n (-0.95) and θ_r and n (0.95). On the one hand, these results show that all the parameters are well estimated since, as compared to the prior intervals (given in Table 1), the confidence intervals of the estimated parameters (plotted in Figs. 9-14) are strongly reduced, especially for the parameters α, n and θ_s. On the other hand, compared to the results of scenario 4 which also considers pressure data, k_s is not as well estimated (the mean value is less close to the reference value and the confidence interval is 27% larger). The mean estimated values for θ_r and n also degraded (less close to the reference solution), although their confidence intervals are similar to those of scenario 4 (Figs. 11, 13). The estimated mean value of the parameter α is similar to that in scenario 4. However, its uncertainty is much larger because the credible interval is 77% larger (Fig. 14). The parameters θ_s and a_l are estimated as well as in scenario 4 (in terms of mean estimated value and credible interval).

The last scenario (scenario 7) is similar to the previous one, but the injection period is reduced to $T_{inj} = 3000\,\mathrm{min}$. The results depicted in Fig. 8 show similar correlations between the parameters as for $T_{inj} = 5000\,\mathrm{min}$. However, a significant improvement is observed for the mean estimated values, which approach the reference solution for k_s, θ_r, n and a_l (Figs. 9, 11, 13, 14). The uncertainties of k_s, α and a_l are also reduced by approximately 40%, 15% and 70%, respectively. The parameter θ_s is estimated as well as in scenario 6. The improvement of the parameter estimation in this last scenario compared to the previous one can be explained by the fact that the injection of water and solute contaminant is stopped once the concentration reaches the column outlet. Hence, the injected volume (0.015x3000 = 45cm^3/cm^2) is slightly less than the pore volume (120x0.43=51 cm^3/cm^2). Thus, when the injection is stopped, the column is not fully saturated and the outlet flux strongly reduces (see the asymptotic behavior of the cumulative outflow when the injection is stopped in Fig. 1c). As a consequence, the concentration profile increases smoothly (see Fig. 1f) until reaching its maximum value in contrast to the sharp front observed for $T_{inj} = 5000\,\text{min}$ in the scenario 6 (see Fig. 1e). Hence, the breakthrough curve obtained with $T_{inj} = 3000\,\text{min}$ is more affected by the hydraulic parameters than the breakthrough curve obtained with $T_{inj} = 5000\,\text{min}$. This explains why a better estimation of the parameters is observed for the last scenario compared to the scenario 6.

5. Conclusions

In this work, estimation of hydraulic and transport soil parameters have been investigated using synthetic infiltration experiments performed in a column filled with a sandy clay loam soil, which was subjected to continuous flow and solute injection over a period T_{inj}.

The saturated hydraulic conductivity, the saturated and residual water contents, the Mualem-van Genuchten shape parameters and the longitudinal dispersivity are estimated in a Bayesian framework using the Markov Chain Monte Carlo (MCMC) sampler. Parameter estimation is performed for different scenarios of data measurements.

The results reveal the following conclusions:

1. All hydraulic and transport parameters can be appropriately estimated from the described infiltration experiment. However, the accuracy differs and depends on the

type of measurement and the duration of the injection T_{inj}, even if the water content remains close to saturated conditions.

2. The use of concentration measurements at the column outflow, in addition to traditional measured variables (water content, pressure head and cumulative outflow), reduces the hydraulic parameters uncertainties, especially that of the saturated water content (comparison between scenario 2 and scenario 3).
3. The saturated hydraulic conductivity is estimated with the same order of accuracy, independent of the observed variables.
4. The estimation of the dispersivity is sensitive to the injection duration. The scenarios 5 and 7 with $T_{inj} = 3000\,\mathrm{min}$ yield much more accurate dispersivity estimations than scenarios 4 and 6 with $T_{inj} = 5000\,\mathrm{min}$ due to the extended spreading of the solute observed for $T_{inj} = 3000\,\mathrm{min}$.
5. A better identifiability of the soil parameters is obtained using C instead of θ measurements, in combination with h and Q data (comparison between scenario 2 and scenario 4).
6. Using only non-intrusive measurements (cumulative outflow and output concentration) yields satisfactory estimation of all parameters (scenario 7). The uncertainty of the parameters significantly decreases when the injection of water and solute is maintained for a limited period (comparison between scenario 6 and scenario 7).

This last point has practical applications for designing simple experimental setups dedicated to the estimation of hydrodynamic and transport parameters for unsaturated flow in soils. The setup has to be appropriately equipped to measure the cumulative water outflow (e.g., weighing machine) and the solute breakthrough at the column outflow (e.g., flow through electrical conductivity). The injection should be stopped as soon as the solute concentration reaches the outflow. The accuracy of the estimation of θ_r, α and n improves by adding pressure measurements inside the column, close to the injection.

These results are of course related to the models and experimental conditions we used. This work can be extended to different types of soils, water retention and/or relative permeability functions to evaluate the interest of coupling flow and transport for parameter identification. This work can also be extended to reactive solutes.

Acknowledgments

The authors are grateful to the French National Research Agency, which funded this work through the program AAP Blanc - SIMI 6 project RESAIN (n°ANR-12-BS06-0010-02).

References

Ades A.E., G. Lu. 2003. Correlations between parameters in risk models: estimation and propagation of uncertainty by Markov Chain Monte Carlo. Risk Anal. 23(6):1165-72.

Carrera J., and S.P. Neuman. 1986. Estimation of aquifer parameters under transient and steady conditions: 2. Uniqueness, stability and solution algorithms. Water Resour. Res., 22, 211–227.

Durner W., B. Schultze, T. Zurmühl. 1999. State-of-the-art in inverse modeling of inflow/outflow experiments. p661-681. In M.Th. van Genuchten et al. (ed.) Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media, Proc. Int. Worksh. Riverside, CA. Univ. of California, Riverside.

Durner W., S.C. Iden. 2011. Extended multistep outflow method for the accurate determination of soil hydraulic properties near water saturation. Water Resour. Res. 47:W08526. doi: 10.1029/2011WR010632

Dusek J, M. Dohnal, M. Snehota, M. Sobotkova, C. Ray, T. Vogel. 2015. Transport of bromide and pesticides through an undisturbed soil column: a modeling study with global optimization analysis. J Contam Hydrol. Apr-May;175-176:1-16. doi: 10.1016/j.jconhyd.2015.02.002.

Eching S.O., J.W. Hopmans. 1993. Optimization of hydraulic functions from transient outflow and soil water pressure data. Soil Sci. Soc. Am. J. 57:1167-1175. doi:10.2136/sssaj1993.03615995005700050001x

Eching S.O., J.W. Hopmans, O. Wendroth. 1994. Unsaturated Hydraulic Conductivity from Transient Multistep Outflow and Soil Water Pressure Data. Soil Sci. Soc. Am. J. 58: 687-95 doi:10.2136/sssaj1994.03615995005800030008x

Fahs M., A. Younes, F. Lehmann. 2009. An easy and efficient combination of the Mixed Finite Element Method and the Method of Lines for the resolution of Richards'

Equation. Environmental Modelling & Software ;24:1122–1126. doi:10.1016/j.envsoft.2009.02.010

Fahs M., A. Younes, P. Ackerer. 2011. An efficient implementation of the method of lines for multicomponent reactive transport equations. Water air and soil pollution, vol. 215, no1-4, pp. 273-283. doi:10.1007/s11270-010-0477-y

Farthing M.W., C.E. Kees, C.T. Miller. 2003. Mixed finite element methods and higher order temporal approximations for variably saturated groundwater flow. Adv. in Water Resour. 26:373-394. doi: 10.1016/S0309-1708(02)00187-2

Gallagher M., J. Doherty. 2007. Parameter estimation and uncertainty analysis for a watershed model. Environmental Modelling & Software 22, 1000-1020. doi:10.1016/j.envsoft.2006.06.007

Gelman A., J.B. Carlin, H.S. Stren, D.B. Rubin. 1997. Bayesian data analysis, Chapmann and Hall, London.

Gelman A., D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7:457-472.

Hopmans J.W., J. Simunek, N. Romano, W. Durner. 2002. Simultaneous determination of water transmission and retention properties. Inverse Methods. p963-1008. In J.H. Dane and G.C. Topp (ed.) Methods of Soil Analysis. Part 4. Physical Methods. Soil Science Society of America Book Series No. 5.

Hudson D.B., P.J. Wierenga, R.G. Hills 1996. Unsaturated hydraulic properties from upward flow into soil cores. Soil Sci Soc Am J;60:388±96.

Inoue M., J. Šimůnek J.W. Hopmans, and V. Clausnitzer. 1998. In situ estimation of soil hydraulic functions using a multistep soil-water extraction technique. Water Resour. Res. 34:1035–1050.

Inoue M., J. Šimůnek, S. Shiozawa, J.W. Hopmans. 2000. Simultaneous estimation of soil hydraulic and solute transport parameters from transient infiltration experiments, Adv. in Water Resour. 23 (7). Doi : 10.1016/S0309-1708(00)00011-7.

Kahl G.M., Y. Sidorenko, B. Gottesbüren. 2015. Local and global inverse modelling strategies to estimate parameters for pesticide leaching from lysimeter studies. Pest Manag Sci. Apr;71(4):616-31. doi: 10.1002/ps.3914.

Kool J.B., J.C. Parker, M.Th van Genuchten. 1985. Determining soil hydraulic properties from one-step outflow experiments by parameter estimation: I. Theory and numerical studies. Soil Sci Soc Am J;49:1348±54.

Kool J.B., and J.C. Parker. 1988. Analysis of the inverse problem for transient unsaturated flow. Water Resour. Res. 24:817–830.

Laloy E., J.A. Vrugt. 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing, Water Resour. Res., 48, W01526. doi:10.1029/2011WR010608

Li H., M.W. Farthing, C.N. Dawson, C.T. Miller. 2007. Local discontinuous Galerkin approximations to Richards' equation. Adv. in Water Resour. 30:555–575. doi: 10.1016/j.advwatres.2006.04.011

Marquardt DW. 1963. An algorithm for least-squares estimation of nonlinear parameters. SIAM J Appl Math;11:431±41.

Mertens J., G. Kahl, B. Gottesbüren, J. Vanderborght. 2009. Inverse Modeling of Pesticide Leaching in Lysimeters: Local versus Global and Sequential Single-Objective versus Multiobjective Approaches Vadose Zone J. 8(3). doi: 10.2136/vzj2008.0029

Miller C.T., G.A. Williams, C.T. Kelly, M.D. Tocci. 1998. Robust solution of Richards' equation for non uniform porous media. Water Resour. Res. 34:2599–2610. doi: 10.1029/98WR01673

Miller C.T., C. Abhishek, M. Farthing. 2006. A spatially and temporally adaptive solution of Richards' equation. Adv. in Water Resour. 29:525–545. doi: 10.1016/j.advwatres.2005.06.008

Mishra S., J.C. Parker. 1989. Parameter estimation for coupled unsaturated flow and transport. Water Resour Res. 25(3). doi: 10.1029/WR025i003p00385

Mualem Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12:513–522. doi:10.1029/WR012i003p00513

Nasta P., S. Huynh, J.W. Hopmans. 2011. Simplified Multistep Outflow Method to Estimate Unsaturated Hydraulic Functions for Coarse-Textured Soil Sci. Soc. Am. J. 75, p.418. doi:10.2136/sssaj2010.011

Šimůnek, J., and M.Th. van Genuchten. 1997. Estimating unsaturated soil hydraulic properties from multiple tension disc infiltrometer data. Soil Sci. 162:383–398.

Tocci M.D., C.T. Kelly, C.T. Miller. 1997. Accurate and economical solution of the pressure-head form of Richards' equation by the method of lines. Adv. in Water Resour. 20:1–14. doi: 10.1016/S0309-1708(96)00008-5

van Dam J.C., J.N.M. Stricker, P. Droogers. 1992. Inverse method for determining soil hydraulic functions from one-step outflow experiment. Soil Sci Soc Am J. 56:1042±50.

van Dam J.C., J.N.M. Stricker, P. Droogers. 1994. Inverse method to determine soil hydraulic functions from multistep outflow experiments. Soil Sci. Soc. Am. J. 58:647-652. doi:10.2136/sssaj1994.03615995005800030002x

van Genuchten M.Th. 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44(5):892-898. doi:10.2136/sssaj1980.03615995004400050002x

Van Genuchten M.Th. and D.R. Nielsen. 1985. On describing and predicting the hydraulic properties of unsaturated soils. Annales Geophysicae, 1985, 3, 615–628.

Vrugt J.A., W. Bouten. 2002. Validity of first-order approximations to describe parameter uncertainty in soil hydrologic models. Soil. Sci. Soc. Am. J. 66:1740-1751. doi:10.2136/sssaj2002.1740

Vrugt J.A., W. Bouten, H.V. Gupta, J.W. Hopmans. 2003a. Toward improved identifiability of soil hydraulic parameters: On the selection of a suitable parametric model. Vadose Zone J. 2:98–113. doi: 10.2113/2.1.98

Vrugt J.A., H.V. Gupta, W. Bouten, S. Sorooshian. 2003b. A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment for hydrologic model parameters. Water Resour. Res. 39(8):1201, doi:10.1029/2002WR001642.

Vrugt J.A., C.J.F. ter Braak, M.P. Clark, J.M. Hyman, B.A. Robinson. 2008. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. Water Resour. Res., 44, W00B09. doi: 10.1029/2007WR006720.

Younes A., M. Fahs, S. Ahmed. 2009. Solving density driven flow problems with efficient spatial discretizations and higher-order time integration methods. Advances in Water Resources 2009, 32 (3) pp 340-352, doi:10.1016/j.advwatres.2008.11.003.

Younes A., T.A. Mara, N. Fajraoui, F. Lehmann, B. Belfort, H. Beydoun. 2013. Use of Global Sensitivity Analysis to Help Assess Unsaturated Soil Hydraulic Parameters. Vadose Zone J. 12. doi:10.2136/vzj2011.0150.

Zurmühl T. 1996. Evaluation of different boundary conditions for independent determination of hydraulic parameters using outflow methods. In Parameter Identification and Inverse Problems in Hydrology, Geology and Ecology, eds. J.Gottlieb and P. DuChateau. Kluwer, Dordrecht, 1996, pp.165–184.

List of table captions

Table 1. Prior lower and upper bounds of the uncertainty parameters and reference values.

Table 2. Measurement sets and injection periods for the different scenarios. The pressure head h and the water content θ are measured at 5 cm from the top of the column. The cumulative outflow Q and the concentration C are measured at the exit of the column.

Table 3. Summary of the pairwise parameter correlations.

Parameters	Lower bounds	Upper bounds	Reference values
k_s [cm min^{-1}]	0.025	0.1	0.0347
θ_s [-]	0.3	0.5	0.43
θ_r [-]	0.05	0.2	0.09
α [cm^{-1}]	0.01	0.3	0.04
n [-]	1.2	5	1.4
a_l [cm]	0.05	0.6	0.2

Table 1. Prior lower and upper bounds of the uncertainty parameters and reference values.

Scenario	**Measured variables**				**Injection period**	
	h	θ	Q	C	$T_{inj} = 5000\,\mathrm{min}$	$T_{inj} = 3000\,\mathrm{min}$
1	ν		ν		ν	
2	ν	ν	ν		ν	
3	ν	ν	ν	ν	ν	
4	ν		ν	ν	ν	
5	ν		ν	ν		ν
6			ν	ν	ν	
7			ν	ν		ν

Table 2. Measurement sets and injection periods for the different scenarios. The pressure head h and the water content θ are measured at 5 cm from the top of the column. The cumulative outflow Q and the concentration C are measured at the exit of the column.

Scenario					
1	$(k_s, n) = -0.97$	$(k_s, \alpha) = 0.94$			$(\theta_r, \theta_s) = 0.96$
2	$(k_s, n) = -0.98$	$(k_s, \alpha) = 0.94$	$(k_s, \theta_r) = -0.94$	$(\theta_r, n) = 0.98$	
3	$(k_s, n) = -0.97$	$(k_s, \alpha) = 0.91$	$(k_s, \theta_r) = -0.94$	$(\theta_r, n) = 0.99$	
4	$(k_s, n) = -0.98$	$(k_s, \alpha) = 0.95$	$(k_s, \theta_r) = -0.96$	$(\theta_r, n) = 0.99$	
5	$(k_s, n) = -0.96$	$(k_s, \alpha) = 0.93$	$(k_s, \theta_r) = -0.91$	$(\theta_r, n) = 0.98$	
6	$(k_s, n) = -0.95$			$(\theta_r, n) = 0.95$	
7	$(k_s, n) = -0.95$			$(\theta_r, n) = 0.94$	

Table 3. Summary of the pairwise parameter correlations.

List of figure captions

Fig. 1. (a) Pressure head at 5 cm below the soil surface, (b) water content at 5 cm below the soil surface, (c) cumulative outflow, (d) retention curve, (e) output concentration for T_{inj} = 5000 and (f) for T_{inj}= 3000 min. Solid lines represent model outputs and dots represent the sets of perturbed data serving as conditioning information for model calibration.

Fig. 2. MCMC solutions for the transport scenario 1. The diagonal plots represent the inferred posterior probability distribution of the model parameters. The off-diagonal scatterplots represent the pairwise correlations in the MCMC drawing.

Fig. 3. MCMC solutions for transport scenario 2 [see Fig. 2 caption].

Fig. 4. MCMC solutions for transport scenario 3 [see Fig. 2 caption].

Fig. 5. MCMC solutions for transport scenario 4 [see Fig. 2 caption].

Fig. 6. MCMC solutions for transport scenario 5 [see Fig. 2 caption].

Fig. 7. MCMC solutions for transport scenario 6 [see Fig. 2 caption].

Fig. 8. MCMC solutions for transport scenario 7 [see Fig. 2 caption].

Fig. 9. Posterior mean values and 95% confidence intervals of the saturated hydraulic conductivity for the different scenarios.

Fig. 10. Posterior mean values and 95% confidence intervals of the saturated water content for the different scenarios.

Fig. 11. Posterior mean values and 95% confidence intervals of the residual water content for the different scenarios.

Fig. 12. Posterior mean values and 95% confidence intervals of the shape parameter α for the different scenarios.

Fig. 13. Posterior mean values and 95% confidence intervals of the shape parameter n for the different scenarios.

Fig. 14. Posterior mean values and 95% confidence intervals of dispersivity for the different scenarios.

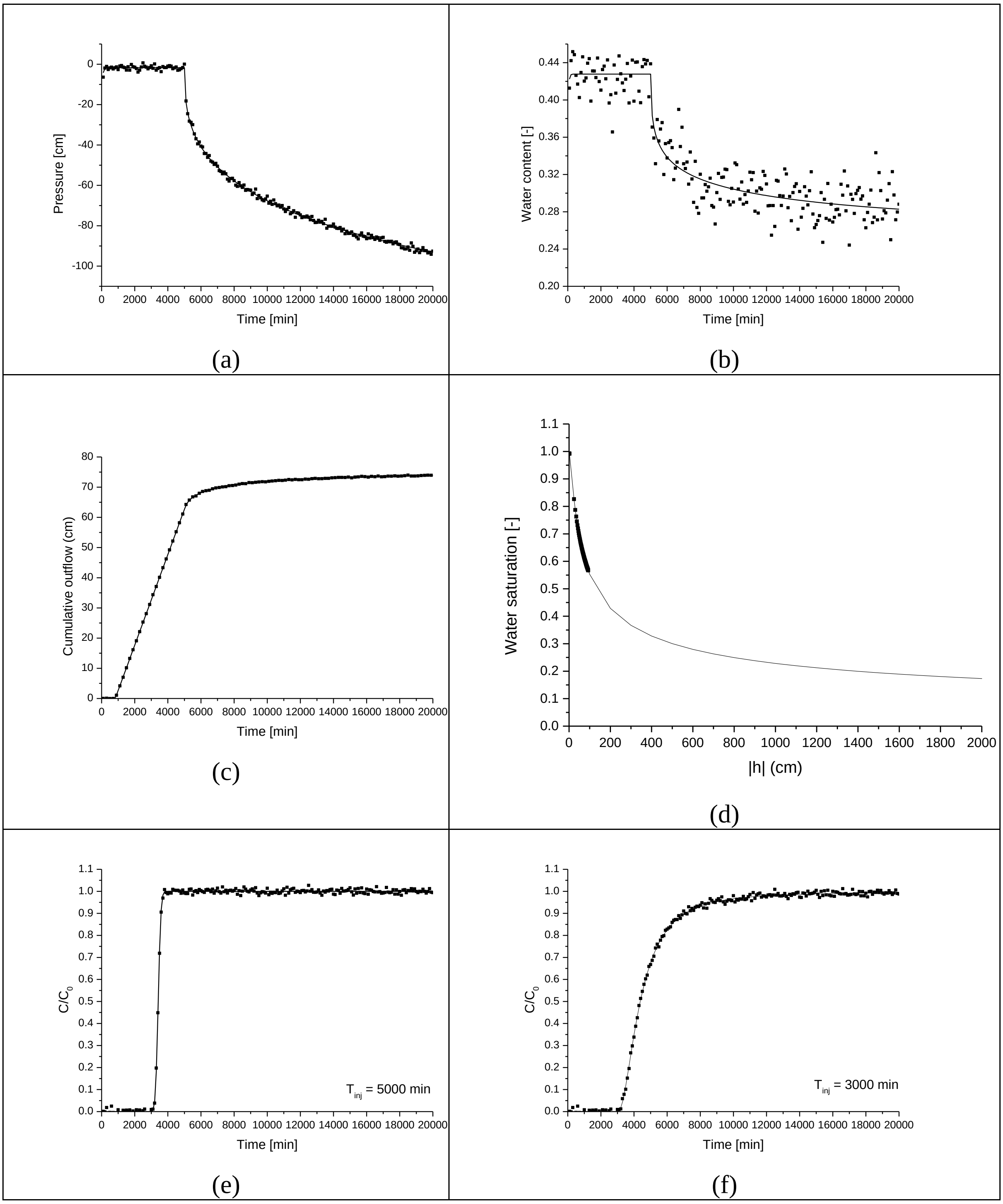

Fig. 1. (a) Pressure head at 5 cm below the soil surface, (b) water content at 5 cm below the soil surface, (c) cumulative outflow, (d) retention curve, (e) output concentration for T_{inj} = 5000 and (f) for T_{inj}= 3000 min. Solid lines represent model outputs and dots represent the sets of perturbed data serving as conditioning information for model calibration.

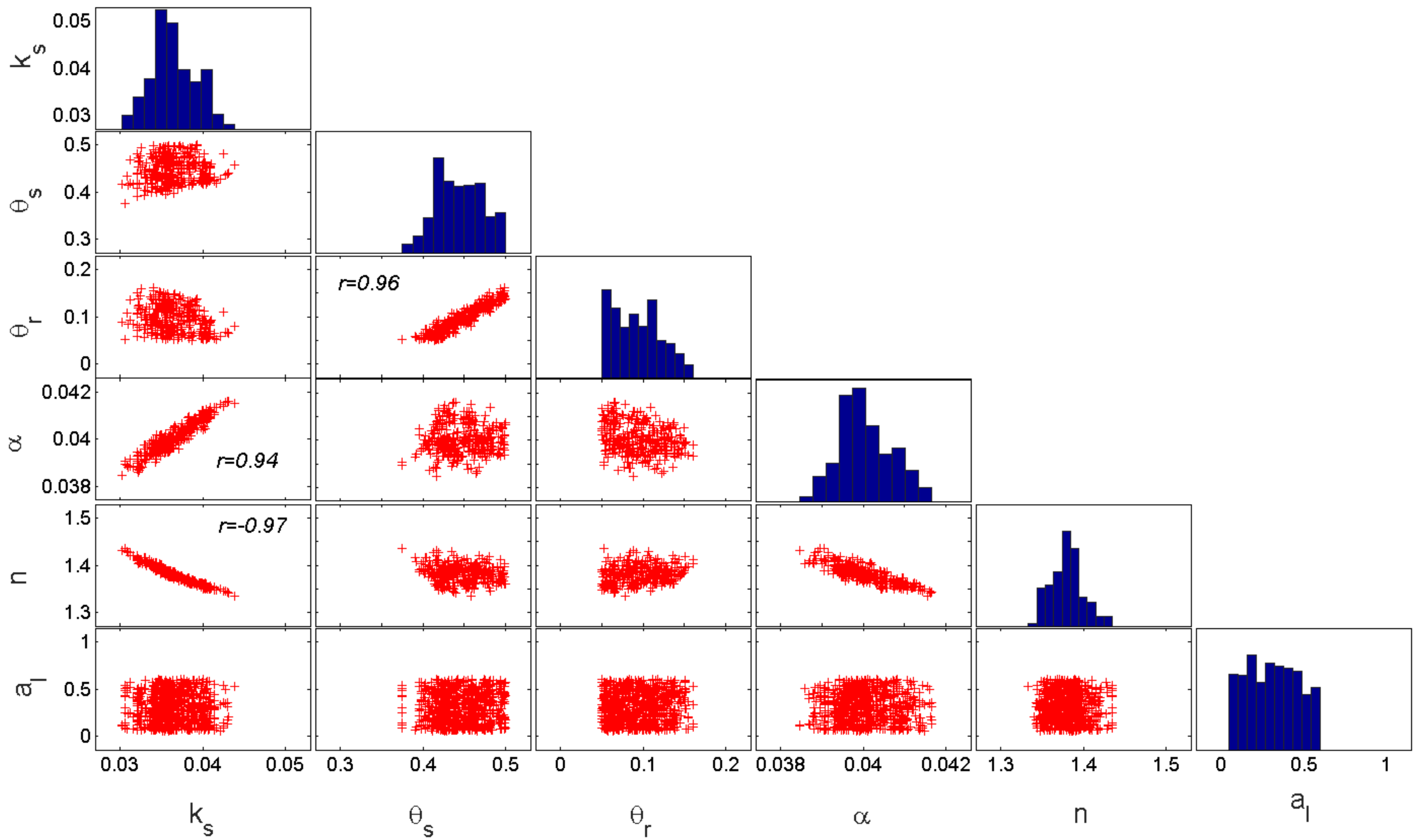

Fig. 2. MCMC solutions for the transport scenario 1. The diagonal plots represent the inferred posterior probability distribution of the model parameters. The off-diagonal scatterplots represent the pairwise correlations *r* in the MCMC draws.

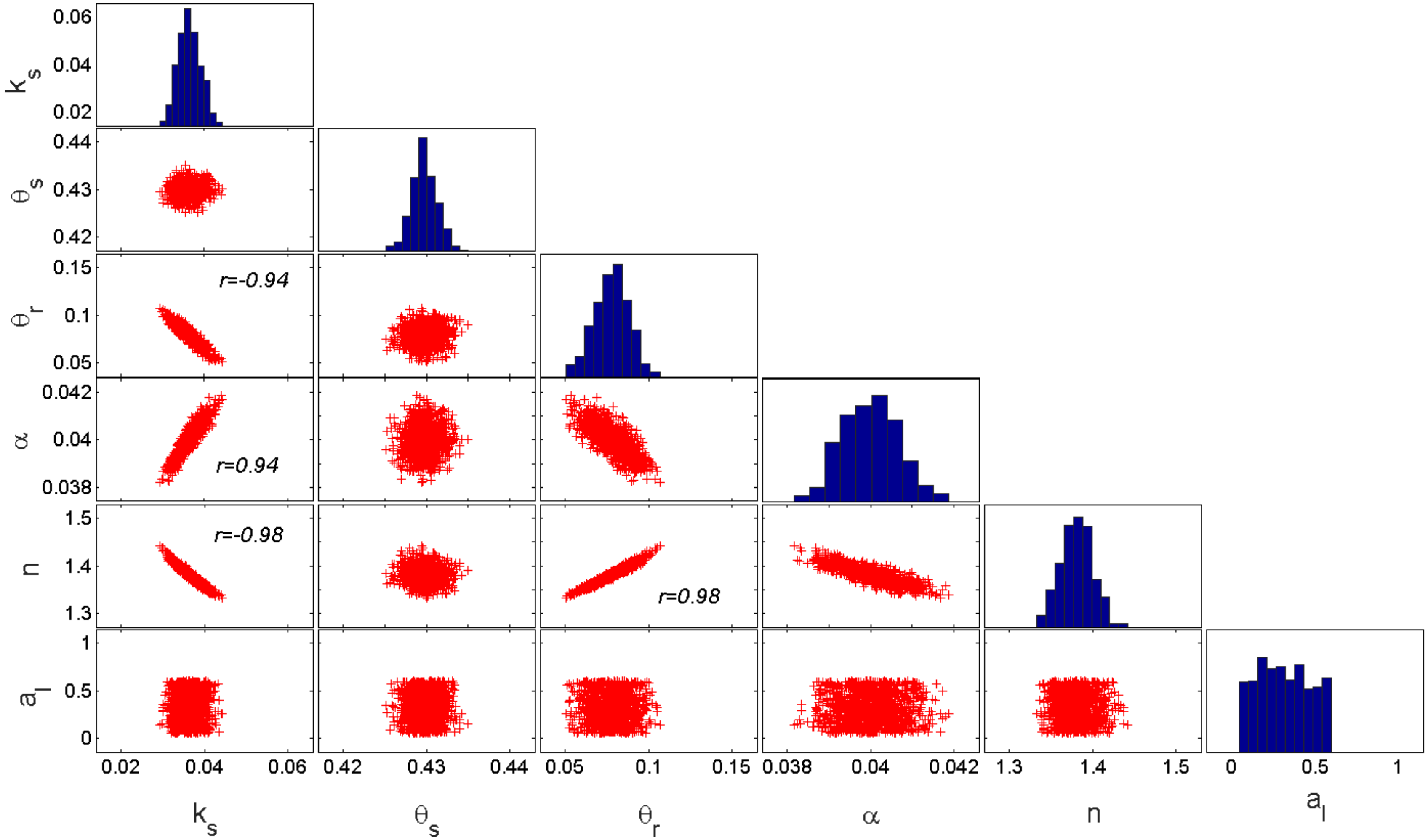

Fig. 3. MCMC solutions for transport scenario 2 [see Fig. 2 caption].

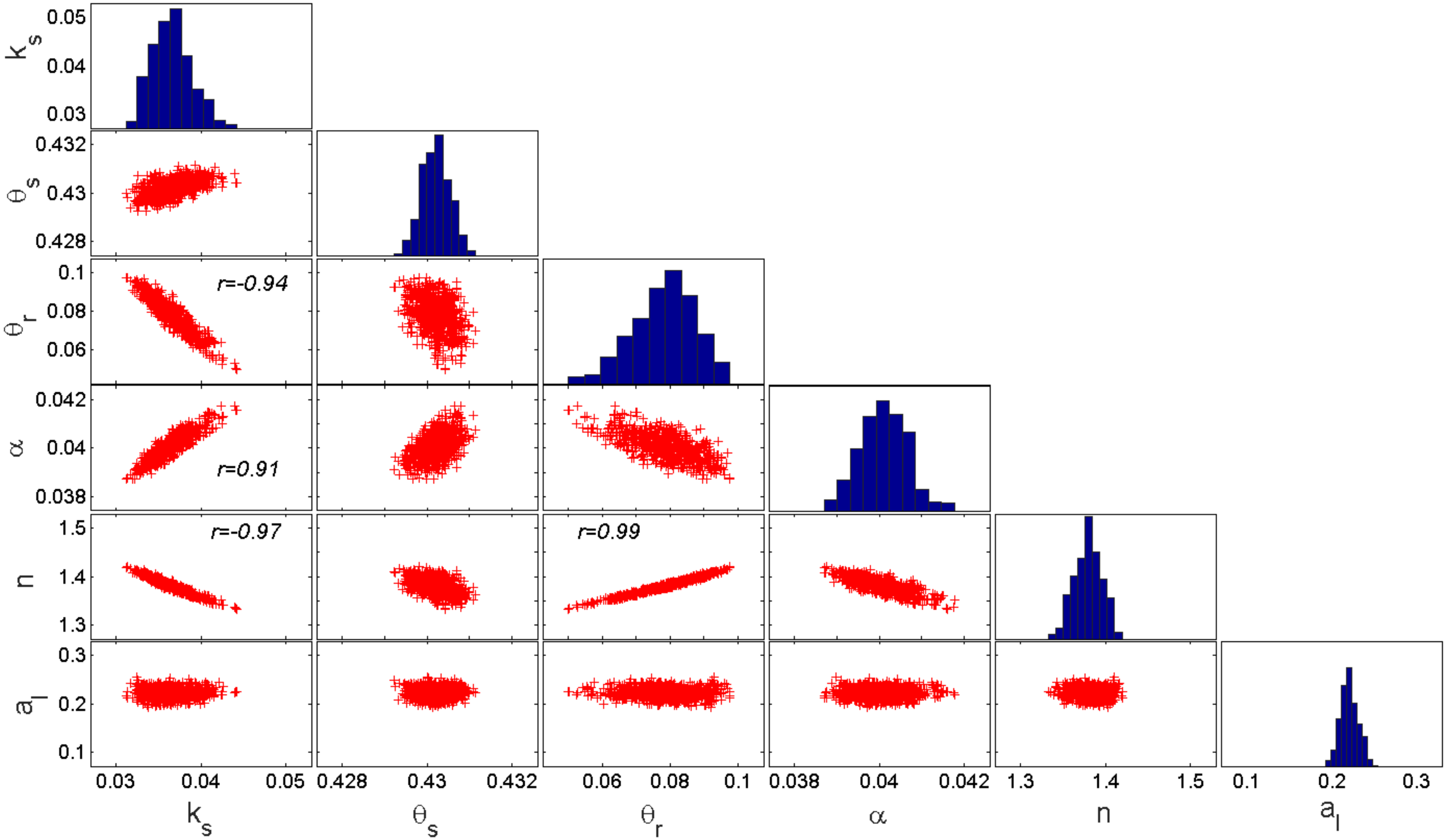

Fig. 4. MCMC solutions for transport scenario 3 [see Fig. 2 caption].

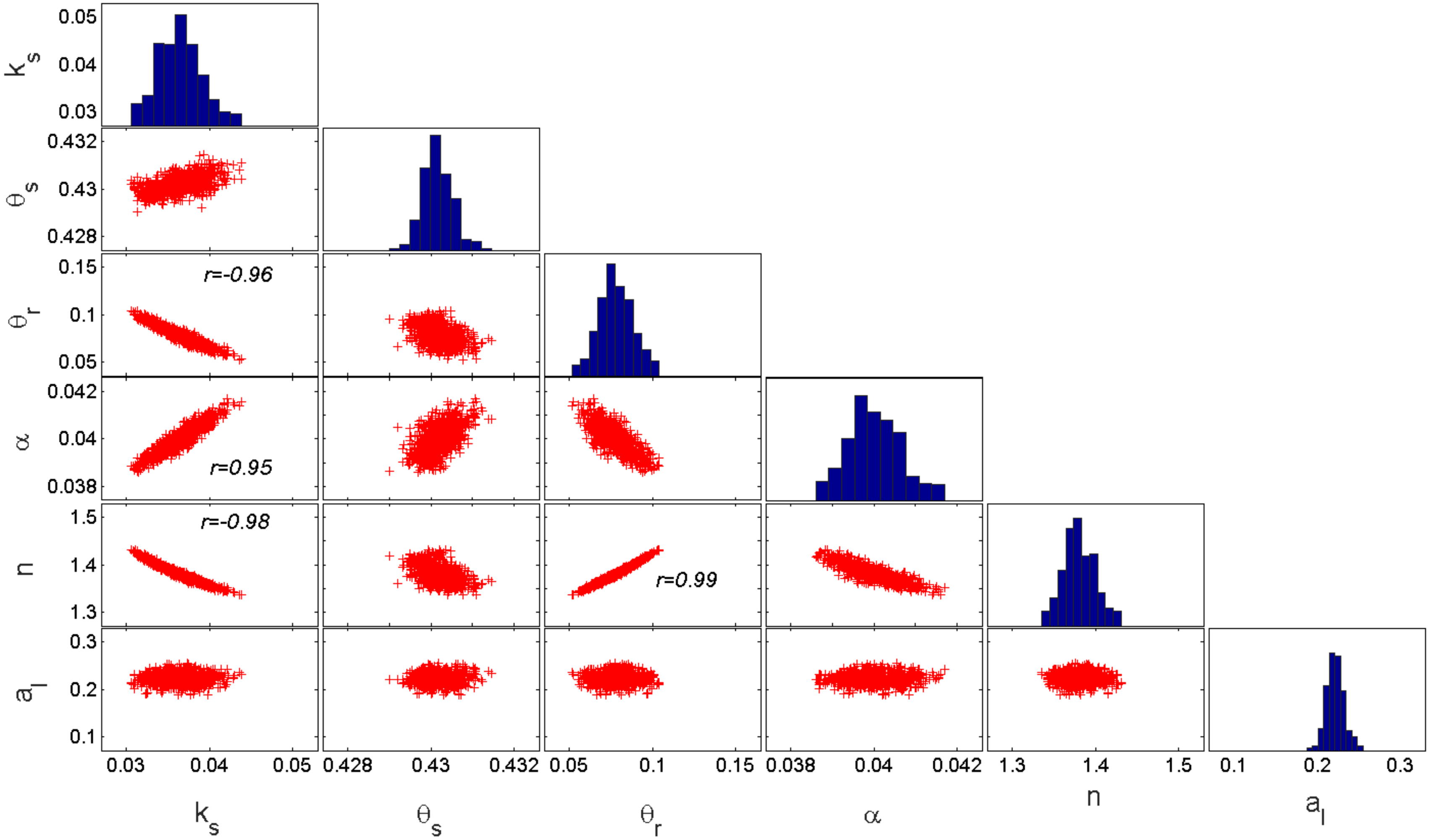

Fig. 5. MCMC solutions for transport scenario 4 [see Fig. 2 caption].

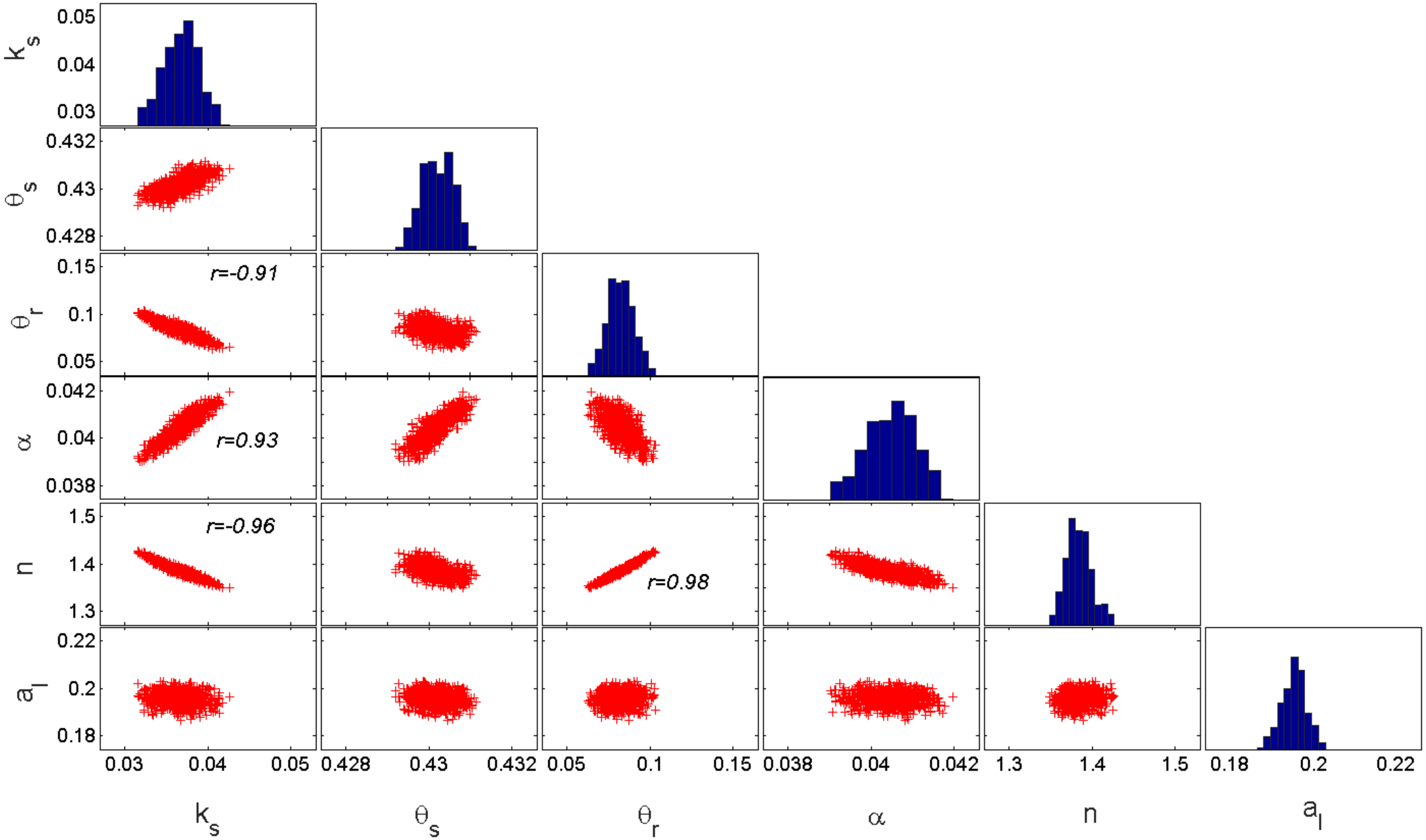

Fig. 6. MCMC solutions for transport scenario 5 [see Fig. 2 caption].

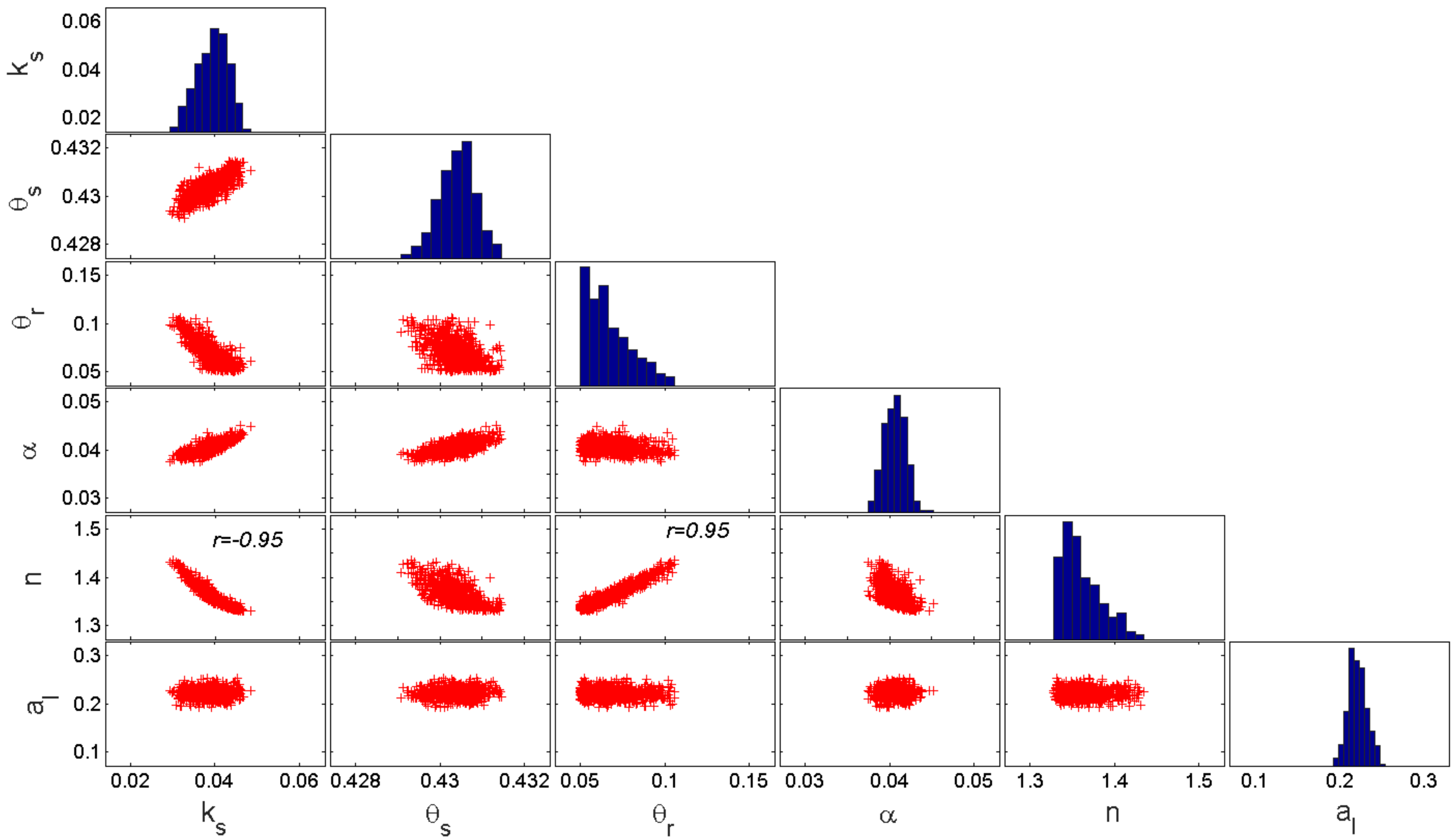

Fig. 7. MCMC solutions for transport scenario 6 [see Fig. 2 caption].

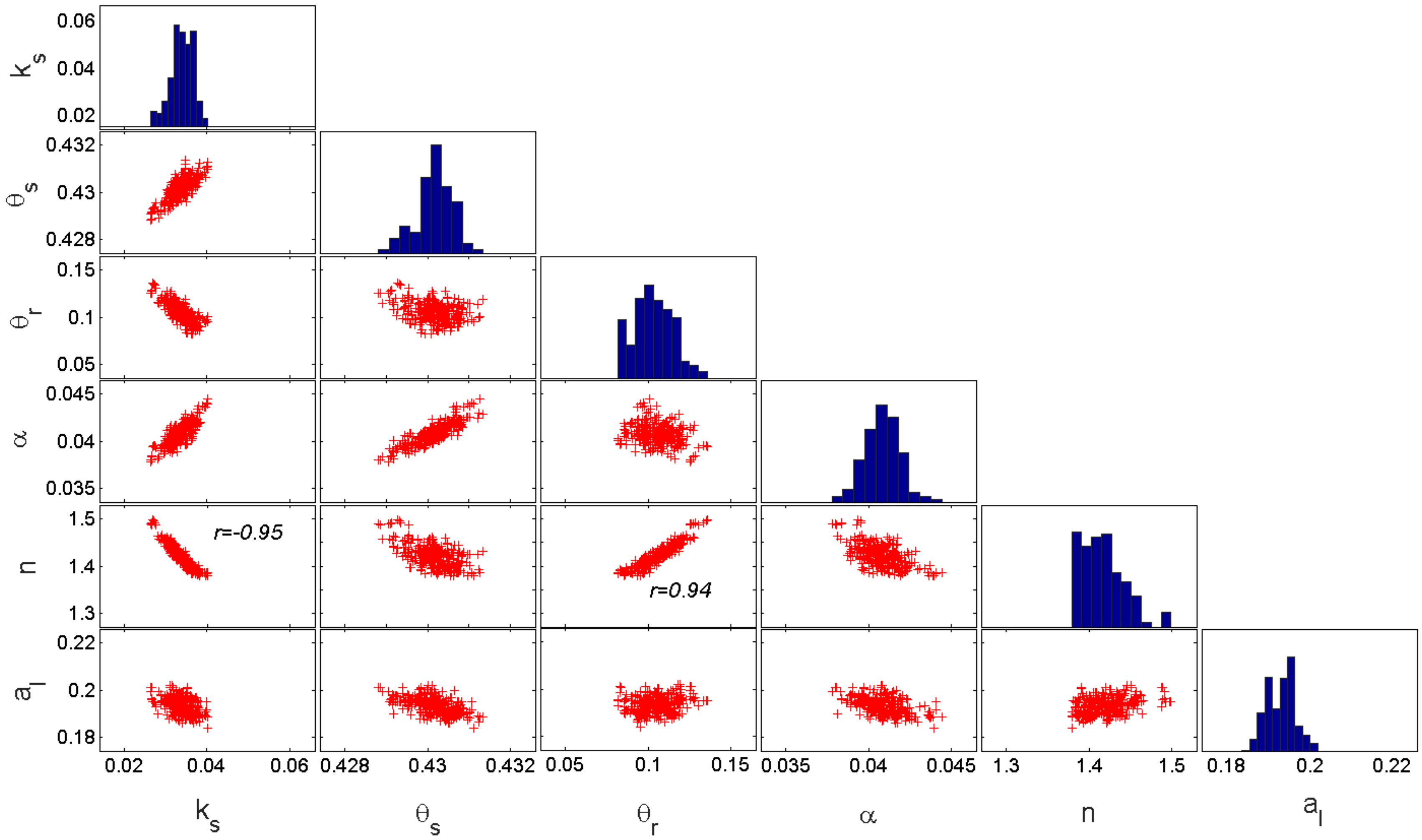

Fig. 8. MCMC solutions for transport scenario 7 [see Fig. 2 caption].

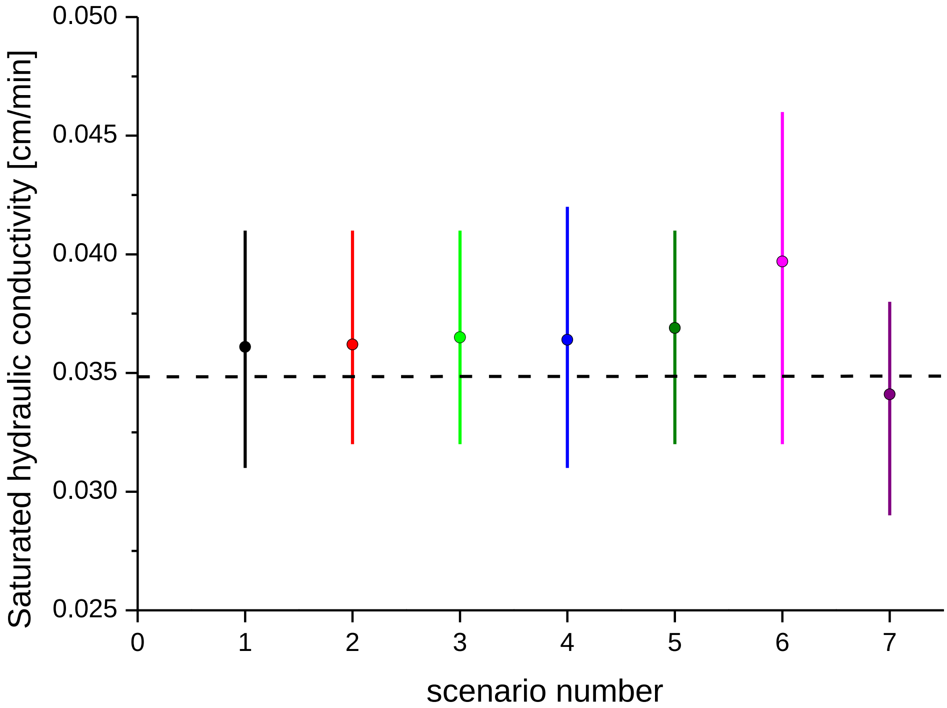

Fig. 9. Posterior mean values and 95% confidence intervals of the saturated hydraulic conductivity for the different scenarios.

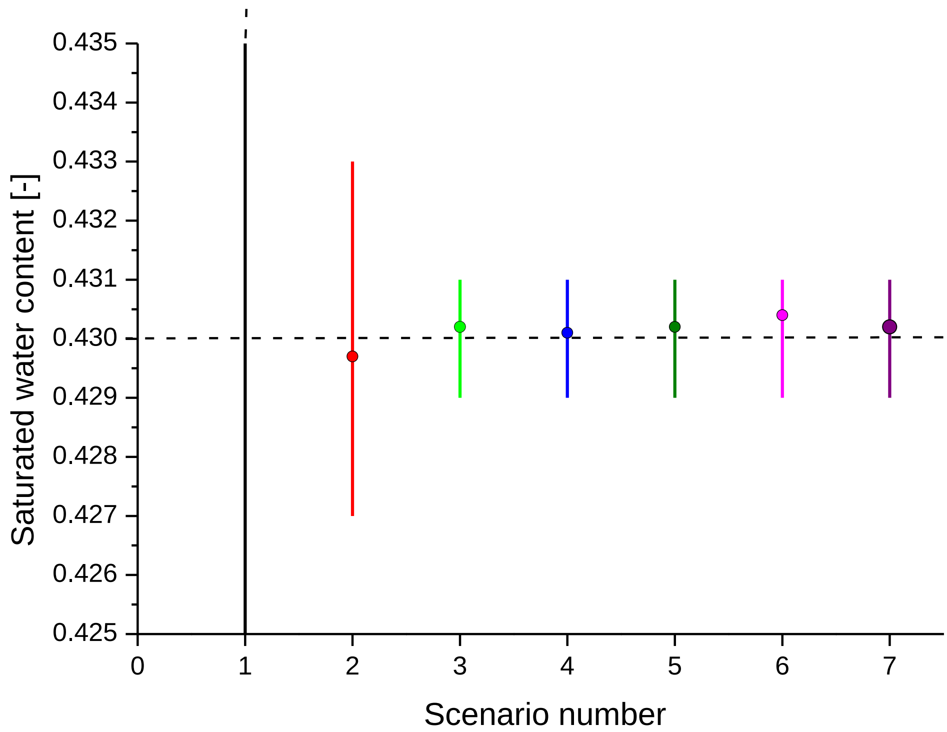

Fig. 10. Posterior mean values and 95% confidence intervals of the saturated water content for the different scenarios.

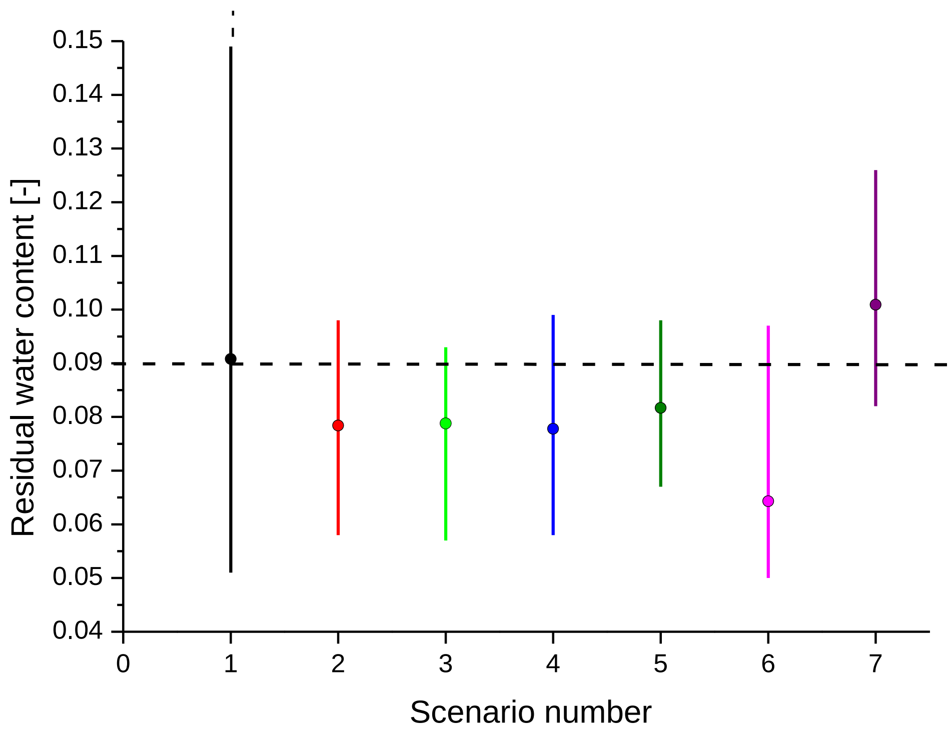

Fig. 11. Posterior mean values and 95% confidence intervals of the residual water content for the different scenarios.

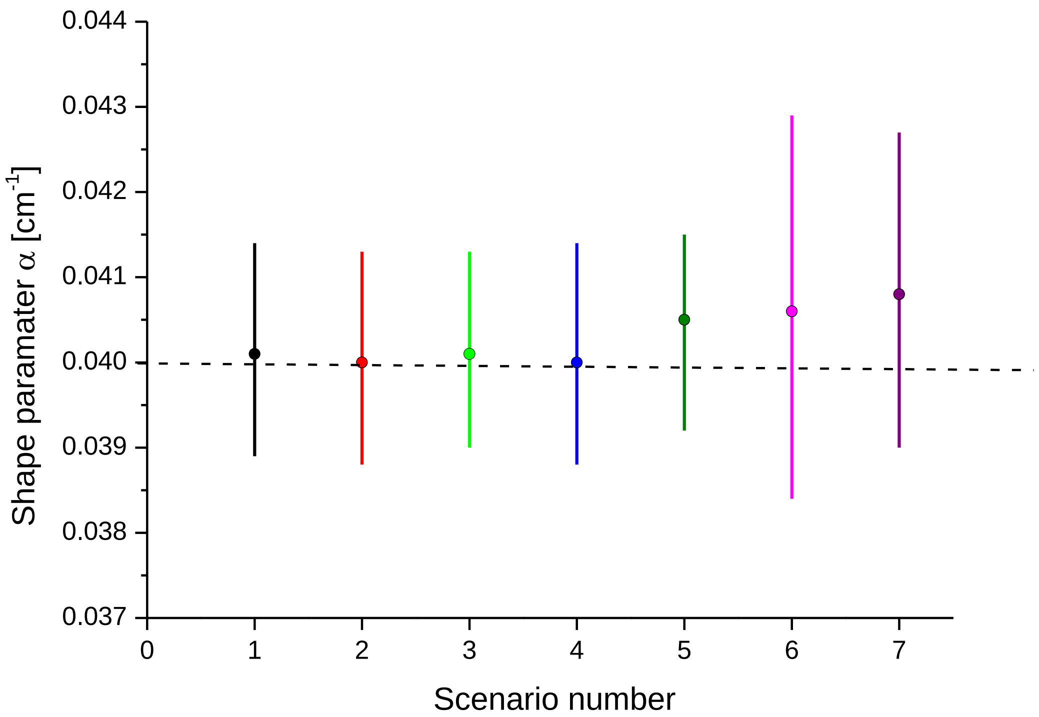

Fig. 12. Posterior mean values and 95% confidence intervals of the shape parameter α for the different scenarios.

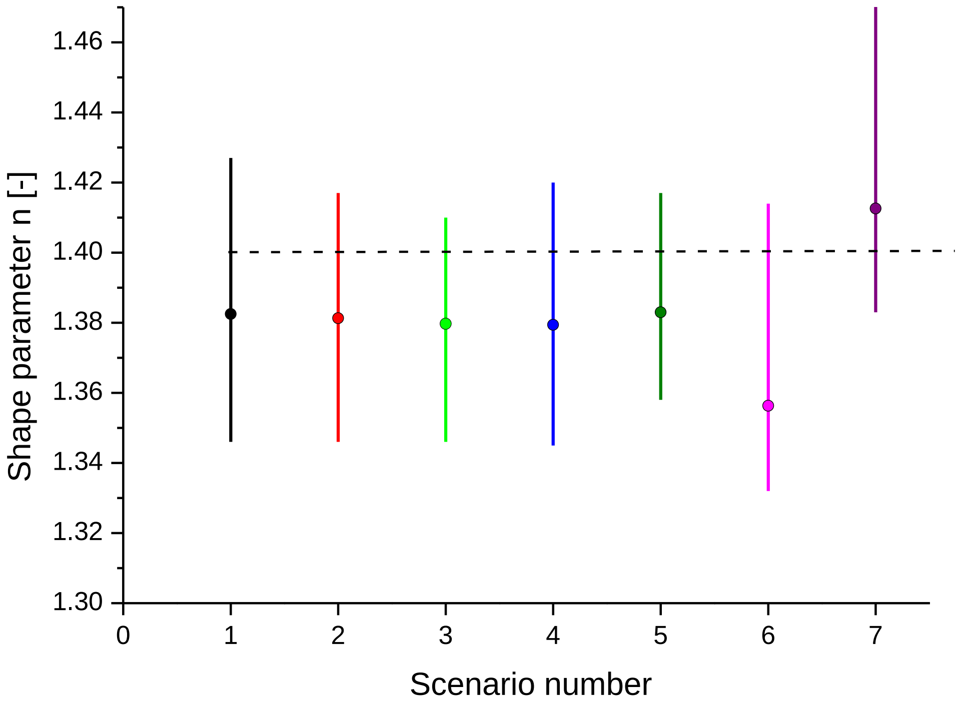

Fig. 13. Posterior mean values and 95% confidence intervals of the shape parameter n for the different scenarios.

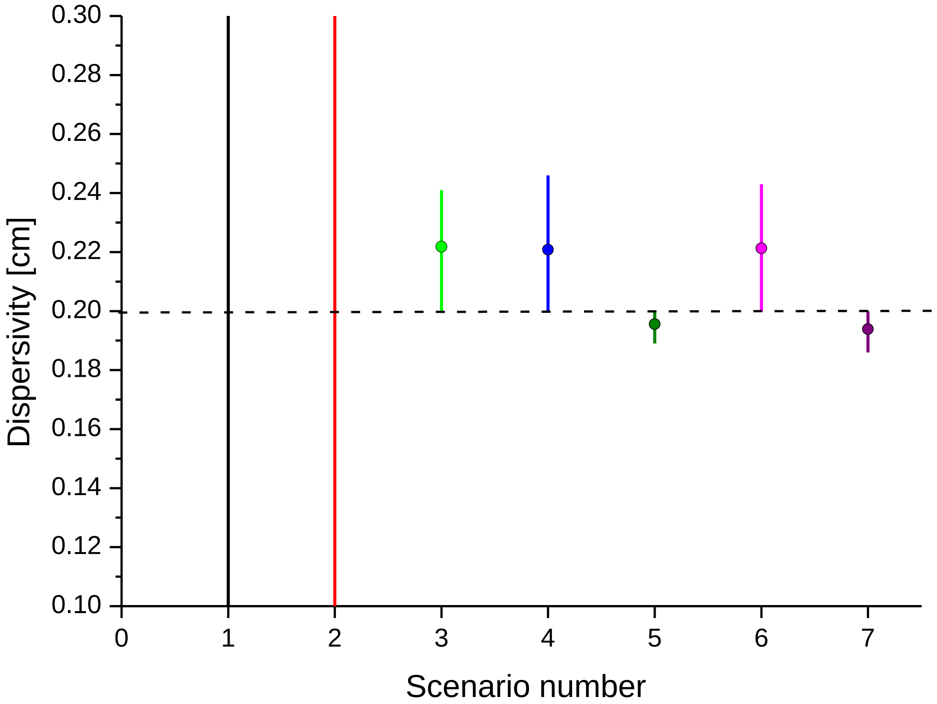

Fig. 14. Posterior mean values and 95% confidence intervals of dispersivity for the different scenarios.